

**Prírodovedecká fakulta, UNIVERZITA KARLOVA, Praha**

**Mapovanie komplexných znakov na modeli kmeni  
konsomických myší**

Bakalárska práca

2007

Školiteľ:

**Prof. Mudr. Jiří Forejt DrSc.**

**Mária Dzúr-Gejdošová**



## Obsah

Abstrakt.....	3
1.Úvod .....	4
2.1. QTL – quantitative trait locus .....	4
2.2. Genetické mapy a markery .....	5
2.3. Selekcia za asistencie markerov.....	6
3. Model pre štúdium QTL.....	6
3.1. Experimentálne kríženia.....	6
3.2. Myš ako modelový organizmus.....	7
3.3. Konsomické kmene myší.....	9
3.4. Heterogénne kmene ako nástroj pre QTL mapovanie.....	9
4. Mapovanie QTL.....	10
4.1. Princíp a postup mapovania QTL.....	10
4.1.1. Detekcia QTL.....	11
4.1.2. Lokalizácia alebo jemné mapovanie.....	12
4.1.3. Sekvenčná analýza.....	12
4.2. Metódy pre analýzu QTL.....	13
4.2.1 Analýza jednotlivých markerov.....	13
4.2.2. Intervalové mapovanie.....	14
4.2.2.1. Metóda maximálnej pravdepodobnosti.....	14
4.2.2.2. Regresná metóda.....	15
4.2.3. Zložené intervalové mapovanie.....	16
4.3. Počítačové programy pre QTL mapovanie.....	16
5. Záver.....	17
Skratky.....	18
Prehľad použitej literatúry.....	19

## **Abstrakt**

Mapovanie lokusov kvantitatívnych znakov (QTL) umožňuje následne identifikovať gény ktoré ich podmieňujú. To vedie k charakteristike a možnej manipulácii s týmito génmi.

V posledných rokoch prinieslo pokrok pre detekciu QTL zostavenie genómových máp s dobre definovanými markermi a k staršej metóde analýzy jednotlivých markerov pribudli aj nové metódy mapovania ako intervalové mapovanie založené na metóde maximálnej pravdepodobnosti či regresnej analýze, a ešte vylepšené zložené intervalové mapovanie. Tiež sa zdokonalili počítačové štatistické programy. Doteraz ale stále bolo objavených len málo génov pre komplexné znaky. Pre urýchlenie hľadania QTL je vhodné použiť kmene konsomických myší. Boli pripravené v našom laboratóriu krížením dvoch inbredných kmeňov myší, PWD/Ph a C57BL/6 tak, že každý kmeň nesie jeden kompletný chromozóm PWD pôvodu na pozadí B6 kmeňa. Zúžujú tak oblasť s predpokladanou prítomnosťou QTL iba na substituovaný chromozóm. Spätným krížením s rodičovským kmeňom (B6) je potom možné ešte zmenšiť oblasť na chromozóme s domnelým QTL.

## **Abstract**

Mapping of quantitative trait loci (QTL) is a prerequisite to identify the genes underlying the traits. This allows subsequent characterizations and possible manipulations of such genes. Recently, a progress has been made in detection of QTLs due to the availability of genome maps with well-defined molecular markers. Besides the older methods, based on single-marker analysis, new mapping methods have been generated, such as interval mapping based on maximum likelihood or regression analysis, and improved composite interval mapping. For all these methods the computational statistics programmes were created. Until now however, only a few genes controlling the complex traits have been discovered. Mouse consomic strains bring up a new potential to speed up the process of finding the QTLs and identifying the responsible genes. In our laboratory, consomic mice were prepared by using the PWD/Ph inbred strain as a chromosome donor, and C57BL/6 providing the genetic background. Thus, every consomic strain carries one complete chromosome of PWD origin on B6 genetic background. The search for a QTL is therefore reduced only to the substituted chromosome. To make the region with a putative QTL much smaller, the backcross analysis of a consomic with C57BL/6 parental strain can be performed.

## 1.Úvod

Tradičné genetické štúdie sa zameriavali na dichotómne znaky a v súvislosti s nimi na prítomnosť či absenciu určitej poruchy, ktoré sú výsledkom mutácie jediného génu (1). Avšak veľké množstvo zaujímavých znakov, vrátane porúch genetickej podstaty ako autizmus, rakovina, diabetes a mnoho iných, je polygénneho charakteru.

Polygénna dedičnosť (tiež dedičnosť kvantitatívnych znakov) sa podľa klasickej teórie uplatňuje u znakov podmienených veľkým množstvom génov, z ktorých sa každý svojím malým účinkom podieľa na výslednom fenotypovom prejave znaku. Okrem samotnej genetickej zložky sa podieľajú na fenotype aj jej interakcie s prostredím. Tieto znaky nazývame aj multigénne alebo geneticky komplexné (2). Môžeme číselne merať veľkosť prejavu kvantitatívnych znakov. Zvyčajne majú normálne rozloženie, ich hodnota sa mení v kontinuálnom gradiente. Oproti monogénnej dedičnosti na polygénnu neplatia klasické pravidlá Mendelovej dedičnosti.

Hľadanie génov pre komplexné znaky je síce obtiažnejšie ako u génov pre monogénne znaky, v posledných rokoch však boli pripravené vhodné modely a zdokonalené metódy pre ich štúdium, a to predstavuje veľkú výzvu pre výskum a objasnenie otázok v oblasti polygénnej dedičnosti.

### 2.1. QTL – quantitative trait locus

QTLs (lokusy kvantitatívnych znakov) sú oblasti genómu obsahujúce gény, ktoré sú zodpovedné za kvantitatívny znak. QTL vyvetľujúce variabilitu jedného znaku sa často môžu nachádzať na viacerých chromozómoch. Môžu byť molekulárne identifikované (napríklad metódou PCR) a pomáhajú tak mapovať regióny genómu obsahujúce gény podieľajúce sa na prejave určitého znaku. Poznanie mnohých QTLs vysvetľujúce fenotypovú variáciu znaku poskytuje predstavu o genetickej architektúre tohto znaku. Tiež nám určí, či je znak kontrolovaný väčším množstvom génov malého efektu alebo menej génmi veľkého efektu.

Obrovský význam pre genetiku komplexných znakov priniesol objav desiatok tisíc intragenových SNPs (single nucleotide polymorphism) a založenie Single Nucleotide Polymorphism Consortium (3). SNP sú rozdiely v DNA sekvenciách genómov dvoch indivíduí rovnakého druhu, keď u každého z jedincov sa na rovnakom lokuse genómu nachádza iný nukleotid. Môžu ukázať, ako vznikli niektoré genetické poruchy a určite sa dajú

významne využiť pri porovnávaní oblasti genómov dvoch príbuzných jedincov, napríklad ak by jeden predstavoval nositeľa určitej poruchy a druhý by bol zdravý.

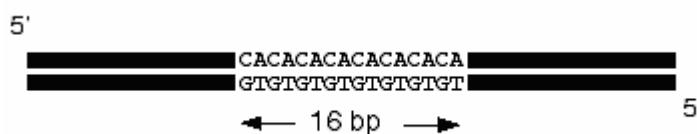
## **2.2. Genetické mapy a markery**

Ako podklad pre vyhľadávanie génov slúžia genetické mapy zostavené na základe dobre charakterizovaných molekulárných markerov (4), skrátene markerov. Pomenovanie marker sa používa všeobecne na popis oblastí s pozorovateľnou variáciou medzi genómami príslušníkov jedného druhu, ktorá je výsledkom mutácie alebo alterácie na rovnakom genetickom lokuse. Obvykle má neutrálny efekt na fenotyp jedinca. Genetické mapy nám poskytujú prehľad najpravdepodobnejšieho usporiadania markerov v genóme pozdĺž chromozómov. Marker môže byť použitý ako orientačný prvok na mape, pokiaľ je daný úsek DNA dedený z rodičov na potomka podľa štandardných pravidiel dedičnosti. Môže sa nachádzať vo vnútri génu kódujúceho znak kvalitatívny, so zreteľnou fyzickou charakteristikou ako napríklad farba očí, i vnútri génu pre komplexný znak, ale môže sa vyskytovať aj v nekódujúcej oblasti génov či úplne mimo oblastí obsahujúcich gény.

Na genetickej mape sú vzdialenosti dvoch bodov či markerov stanovené v závislosti od pravdepodobnosti, s akou medzi nimi dochádza k rekombinácii. Jednotkou vzdialenosti je cM (centiMorgan), ktorý zodpovedá 1 percentu rekombinácie. K rekombinácii, teda výmene úsekov na rovnakých lokusoch môže dôjsť medzi každým párom homologických chromozómov počas meiózy v pohlavných bunkách. Je tým pravdepodobnejšia, čím sú dva úseky na danom chromozóme od seba viac vzdialené. Sú k dispozícii dve metódy na prepočet vzdialeností podľa frekvencií rekombinácií: Haldanovo a Kosambiho mapovanie. Haldanova stratégia (5) sa zameriava na neprítomnosť interferencie, teda že jedno prekríženie neovplyvní výskyt ďalších v neďalekom okolí. Kosambiho naopak uvažuje, že jedno prekríženie zníži pravdepodobnosť výskytu ďalších v blízkosti, a teda predpokladá sa tam negatívna interferencia (6).

Výskyt očakávaných rekombinácií závisí ešte od oblasti genómu ako i na génoch v iných regiónoch genómu. Výrazná odlišnosť je i v genetických mapách človeka u mužov a žien (7). Markery sa dnes identifikujú pomocou molekulárných techník v relatívne krátkom čase (8) a bývajú navrhnuté tak, aby od seba boli vo vzdialenosti okolo 10 cM a aby každý chromozóm lineárne pokrývala sada markerov.

S veľkou výhodou sa ako markery využívajú tzv. mikrosatelity. Sú to oblasti genómu nesúce zoskupené tandemové repetície krátkych sekvencií o dĺžke 2 až 4 nukleotidov (9). Veľmi často sa jedná o desaťnásobné až stonásobné opakovania dinukleotidovej sekvencie CA. V týchto oblastiach často dochádza k ukľznutiu polymerázy a tiež sa tam uplatňuje mechanizmus nerovnomernej rekombinácie, v dôsledku čoho sa mení počet repetícií v mikrosatelite. Vznikajú tak početné alelové možnosti pre jeden mikrosatelit líšiace sa počtom opakovaní. Možno ich amplifikovať metódou PCR a potom elektroforézou rozlíšiť. Jedna metóda nám tak poskytuje možnosť genotypovať za konštantných podmienok stovky až tisíce markerov. Mikrosatelitné markery nám dovoľujú určiť príbuznosť k určitému druhu.



**Obr. 1:** Príklad pre mikrosatelit s najčastejšou repetíciou CA

### 2.3. Selekcia za asistencie markerov

Rozvoj a využitie markerov pri analýze QTL rozšírilo možnosti dedičnosti kvantitatívnych znakov(10). Umožnilo tiež kríženie za asistencie markerov, ktoré prednostne vyberá jedincov s určitými genotypovými variantami ako rodičov pre nasledujúcu generáciu (11). Takto môžeme zmenšiť veľkosť populácie pre ďalšie kríženie a teda i skrátiť čas celého pokusu.

## 3. Model pre štúdium QTL

### 3.1. Experimentálne kríženia

Pre štúdium kvantitatívnych znakov je treba pripraviť, resp. zohnať si dobrú modelovú populáciu. Klasicky sa pripravuje krížením dvoch vybraných geneticky odlišných inbredných kmeňov líšiacich sa navzájom vo fenotypoch. Jedinci sú v rámci svojho kmeňa vybraný tak, aby boli homozygotný pre všetky markery a QTL. Markery sú známe z predchádzajúcich štúdií a sú zostavené pozdĺž celých chromozómov v rovnakých vzdialenostiach (12).

Ak dva inbredné kmene vykazujú v nejakom znaku zjavné fenotypové rozdiely napriek vplyvu prostredia, uvažuje sa, že tento rozdiel má genetický základ. Identita génov zodpovedná za fenotypové rozdiely môže byť odhalená po sérii súrodeneckých krížení dvoch rozdielnych inbredných experimentálnych kmeňov organizmov, začínajúcim rozmnožovaním rodičovskej populácie. Potomkovia 1. filiálnej generácie (F1) potom obsahujú vždy jeden chromozóm v páre od jedného rodiča a druhý chromozóm od rodiča druhého kmeňa. V ďalšom kroku sa krížia jedinci F1 generácie buď medzi sebou, alebo s jedným z rodičovských kmeňov a dostávame generáciu spätných krížencov, z ktorých každý obdržal jeden chromozóm v páre kompletný z inbredného kmeňa, druhý chromozóm od rodiča F1 generácie je však v dôsledku rekombinácie počas meiózy mozaikou chromozómov oboch inbredných kmeňov (1). Spätné kríženie (backcross) je vďaka svojej jednoduchosti pre následné detekovanie QTL najčastejšie využívané. Spätný kríženci majú veľkú výhodu v tom, že na každom lokuse v genóme môžu mať len jeden z dvoch možných genotypov. Produkuje sa mnoho spätných krížencov, väčšinou 100. U každého potomka je potom genotypovaním zistené, ktoré z dvoch možných markerov od rodičov získal. Vyhodnotia sa rekombinačné hodnoty medzi všetkými dvojicami markerov (13). Pre každého potomka sa ešte určí číselná fenotypová hodnota sledovaného znaku. Rodičovské kmene sa v hodnote tohto znaku výrazne líšia. Zdá sa, že jedinci F1 generácie by mali mať priemernú hodnotu v tomto znaku, ale v mnohých prípadoch to nepozorujeme, dosahujú napríklad oveľa vyššie hodnoty oproti priemeru. Zapríčiňuje to práve súbežné posobenie niekoľkých génov rozložených na viacerých lokusoch, alebo dôsledok dominancie resp. recesivity daného markeru.

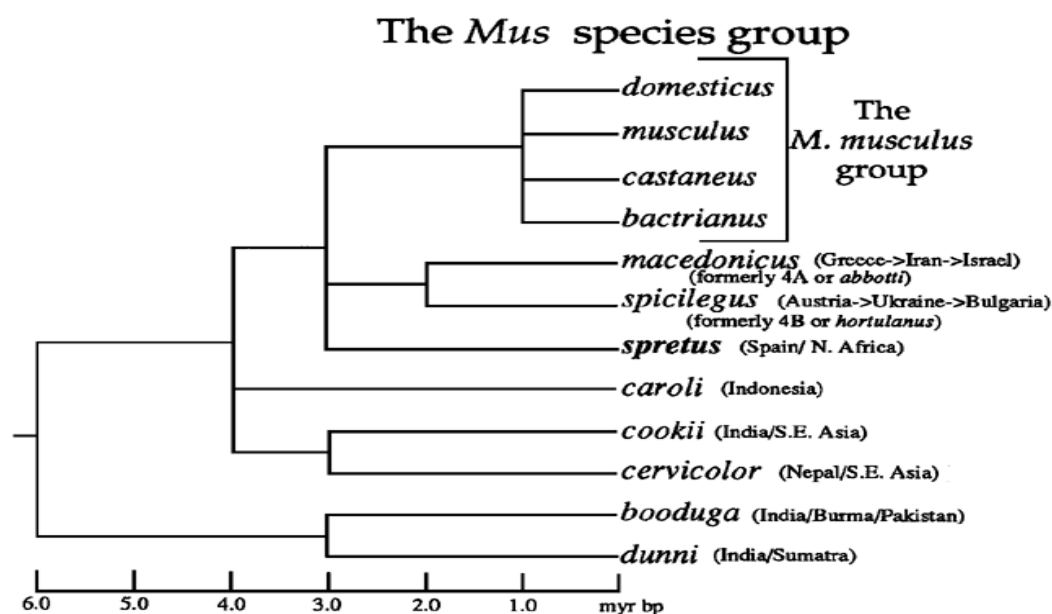
### **3.2. Myš ako modelový organizmus**

Snahy o izoláciu génov zodpovedných za kvantitatívne fenotypy u človeka dosahujú malé úspechy, zväčša kvôli malému účinku QTLs na fenotypovú odlišnosť. Navyše štúdia s ľudskými rodinami nemôžu poskytnúť dostatočné rozlíšenie pre pozičné klonovanie, jedine že by boli rodiny veľmi veľké, a k tomu by ešte výsledky mohli byť narušené vplyvom enviromentálnych či genetických rozdielov medzi sledovanými jedincami a kontrolami (13). Tieto ťažkosti obmedzujú mapovanie QTLs u človeka, a to viedlo k presunu záujmu štúdia na iný živočíšny model, predovšetkým myši. V poslednom storočí sa práve oni stali hlavným cicavčím modelovým organizmom pre genetické výskumy. Sú druhým cicavcom po človeku,



u ktorého bola prečítaná genómová sekvencia (14). Bola zostavená ich genómová mapa (15) pozostávajúca z viacerých identifikovateľných markerov v genome v známych polohách(7). Dobre známy genóm ale i blízka genetická a fyziologická podobnosť človeku predurčila myš na obľúbený model pre mnohých vedcov. Prirodzene sa u nich vyskytuje ako i u iných cicavcov mnoho známych chorôb ako rakovina, osteoporóza, hypertenzia, ale genetickými manipuláciami je u nich možné vyvolať i poruchy, ktoré ich normálne nepostihujú. Vďaka genetickej príbuznosti s človekom tiež môžeme očakávať, že určitý faktor kvantitatívneho znaku u myši bude homologický u človeka.(1) Myš má i ďalšie obrovské výhody: kríženie je kontrolované experimentátorom, a vďaka ich krátkej generačnej dobe, 9 týždňov, je možné pripraviť si požadované jedince v relatívne krátkom čase. Sú navrhnuté i markery mnohých kmeňov a možno ich zohnať v databáze Mouse Simple Sequence Length Polymorphism Database (16).

Genómy väčšiny klasických laboratórnych inbredných myší boli vytvorené hlavne z poddruhov *Mus musculus domesticus* a z 20% z poddruhov *Mus musculus musculus*. Oba tieto poddruhy sa separovane vyvíjali 350 000 rokov, takže u nich prebiehala evolúcia odlišne. To viedlo k akumulácii veľkého množstva SNPs do ich genómu a tým i vzniku množstva dedičných fenotypových variant. Dnes už existujú viaceré genetické analýzy klasických laboratórnych kmeňov a kmeňov odvodených z *M. m. Musculus* a poskytujú mnoho informácií, ktoré uľahčujú ďalšie štúdium .



Obr. č.2: Fylogenetický strom druhov rodu *Mus*.( L. Silver, 1995)

### 3.3. Konsomické kmene myší

Pre štúdium QTL je potrebné nájsť vhodný model, za ktorý si môžeme zvoliť konsomické kmene myší produkované v našom laboratóriu. Boli pripravené krížením dvoch inbredných kmeňov: PWD/Ph pochádzajúceho z *M.m. musculus* a C57/Bl6. Série cielených krížení a spätných krížení za asistencie markerov po niekoľkých generáciách viedli k utvoreniu 21 kmeňov s chromozómovými substitúciami, kde vždy jeden kmeň má jeden chromozóm pôvodom z PWD na genetickom pozadí B6 kmeňa. Kmene označujeme priradením čísla, podľa toho, ktorý substituovaný chromozóm nesú jeho príslušníci. Pri príprave konsomikov nesúcich donorové autozómy alebo X chromozóm sú generácie potomkov spätných krížení kontrolované genotypovaním pre uistenie, že nedošlo k rekombináciám medzi homológnyimi chromozómami hostiteľa a donora. Chromozóm Y je prenesený na pozadie nového kmeňa bez genotypovania. Konsomické kmene urýchľujú hľadanie QTLs v genóme, totiž ak niektorý kmeň vykazuje značnú fenotypovú odlišnosť v určitom znaku, možno to prisúdiť pôsobeniu genotypov na chromozóme, ktorý je v kmeni substituovaný (17). Spätným krížením s jedincami východného inbredného kmeňa, ktorý tvorí genetické pozadie pre substituovaný chromozóm, dostaneme po niekoľkých generáciách potomkov už len so segmentom z pôvodného substituovaného chromozómu. Uvažujeme, že jedinci, ktorí stále vykazujú fenotypovú odlišnosť, nesú segment obsahujúci QTL ktorý vysvetľuje premenlivosť sledovaného znaku. Takto v relatívne krátkom čase definujeme chromozómový región pre hľadaný QTL.

Dokončenie zakladania týchto konsomických kmeňov viedlo k príprave a charakterizovaniu prvej genómovej BAC knižnice myši *Mus m. musculus* (PWD/Ph) (18). BAC (Bacterial Artificial Chromosome) knižnica je zdrojom klonov rozľahlých genomových oblastí, ktoré možno využiť pri sekvenácii či izolácii určitého skúmaného génu.

### 3.4. Heterogénne kmene ako nástroj pre QTL mapovanie

Experimentálne kríženia inbredných kmeňov priniesli síce úspech v mapovaní QTL u mnohých rôznych fenotypov, ale pokusy o jemné mapovanie QTL často stroskotali na objave, že jediný QTL veľkého efektu je v skutočnosti tvorený mnohonásobnými lokusmi s malými efektmi umiestnenými v rovnakej oblasti chromozómu (19). Ďalším potenciálnym problémom

pri detekcii QTL u inbredných kmeňov je značná redukcia genetickej heterogenity v porovnaní s celkovou genetickou variáciou v populácii.

V snahe obísť tieto problémy niektoré štúdie používajú geneticky heterogénne kmene (heterogenous stock, HS) myši vytvorené zo známych predkov. Boli vytvorené krížením z 8 inbredných línií myši (3). Oproti inbredným kmeňom majú HS línie omnoho menej alelových možností než ich zakladateľské kmene a vďaka tomu sa mapuje komplexnejšie.

## **4. Mapovanie QTL**

### **4.1. Princíp a postup mapovania QTL**

V minulosti bola najväčším problémom pri mapovaní nedostupnosť vhodných markerov (20). V posledných rokoch ale boli pre výskum pripravené takmer nevyčerpatelne zásoby molekulárnych markerov a to umožnilo i zostavenie hustejšie pokrytých genetických máp (21). Taktiež boli vylepšené štatistické metódy a počítačové programy pre sledovanie prepojenia markerov a QTL (22).

Pri genetickom mapovaní sú zoradené markery, geny alebo QTL vo vzájomných vzdialenostiach a v závislosti od percenta rekombinácie sú usporiadané do spoločných skupín (23). Mapovanie lokusov kvantitatívnych znakov spočíva v jednoduchej myšlienke, že markery, ktoré sa prenášajú zároveň so zmenou určitého znaku, pravdepodobne budú prepojené s génmi určujúcimi tento znak (24). Pokiaľ nám markery vykazujú výraznú asociáciu, predpokladáme, že budú na genetickej mape umiestnené tesnejšie pri sebe.

Základný princíp je procesmi kríženia oddelovať od seba QTLs a markerové lokusy ktoré sa nerovnomerne prenášajú z generácie na generáciu a hľadá sa štatisticky významná spojitosť medzi segregujúcimi markerovými variantami (genotypmi) a hodnotami znaku (fenotypmi). Pomocou získaných poznatkov súčasne o dvoch či viacerých prepojených markeroch sa dá určiť QTL efekt a jeho pozícia na mape.

Iný postup umožňuje odhadnúť efekt prepojených QTL na základe frekvenčných zmien markerových alel na ich lokuse po niekoľkých generáciách umelo selektovaných podľa sledovaného znaku (25). Vyžaduje ale, aby frekvencie alel v prvej populácii boli známe a jedinci musia byť v stave kompletnej väzbovej nerovnováhy resp. nerovnováhy prepojenia. Metóda väzbovej nerovnováhy vychádza z myšlienky, že medzi niektorými alelami existuje genetická väzba, a teda niektoré kombinácie alel resp. markerov sa objavujú častejšie či

vzácnnejšie než sa predpokladá na základe ich frekvencií v populácii pri náhodnom zostavovaní .

Existujú dva všeobecné prístupy k výberu podkladov pre mapovanie. Jeden sa pokúša preniesť jediný QTL na genetické pozadie iného inbredného kmeňa, pričom pri kríženiach sa v dôsledku rekombinácií zmenšuje segment chromozómu nesúci QTL. Druhý prístup sa snaží zvýšením počtu potomkov F2 generácie detekovať viacero QTL súčasne na základe veľkého množstva rekombinácií v genóme (26).

Predpokladalo sa, že analýza lokusov kvantitatívnych znakov bude rovnako ľahko uskutočniteľná ako u Mendelových znakov (27). Mapovanie monogénnych faktorov zahŕňa tvorbu malého genetického intervalu, kde už jediná rekombinácia zmenší jeho rozsah pod 1 cM vďaka silnému prejavu genotypu vo fenotype. Požadovanú sekvenciu je tak možné nájsť z malého počtu kandidátnych génov. Mapovanie komplexných znakov je ale zložitejšie a podlieha viacerým kritériam. Jeho postup možno rozdeliť do troch základných krokov: detekcia , lokalizácia či jemné mapovanie a sekvenčná analýza. Každý krok vyžaduje rôzne postupy a metódy.

#### **4.1.1. Detekcia QTL**

Najprv musí byť objavené štatisticky významné prepojenie QTL s určitým markerom (28). Pre znaky s normálnym rozdelením je možné požiť klasické parametrické testy, ale keďže nie všetky znaky majú normálne rozdelenie, často sa uchýľuje k prísnejšiemu neparametrickému testu Landera a Kruglyaka (29). Ako model pre detekciu je dobré použiť nejakú segregujúcu populáciu príbuzných jedincov, predovšetkým generácie spätných krížencov spomínaných vyššie. Prvotná detekcia nám vymedzí rozľahlejšiu oblasť v genóme, ktorá bude hľadaný QTL obsahovať. S ohľadom na vlastnosti komplexných znakov sa minimálne intervaly s QTL zvyknú stanovovať na rozpätie aspoň 10 až 30 cM. Veľkosť takého intervalu zaberá u človeka asi 20 až 60 Mb DNA, čo je viac ako 100 génov. Samozrejme by bolo príliš náročné vyhodnocovať funkcie každého z tohto veľkého počtu génov, preto sa v ďalších krokoch pristupuje k presnejšiemu mapovaniu (28).

#### 4.1.2. Lokalizácia alebo jemné mapovanie

Je teda potrebné vypilovať široký interval zahŕňajúci QTL na menší rozsah, čo najviac to bude možné. Preto sú žiadúce ďalšie kríženia s vysokým rozlíšením, k čomu sa využívajú napríklad kongénne kmene. Niekoľko generácií párení umožní priebeh väčšieho počtu meióz a teda i rozchod oblastí v rámci pôvodne širokého intervalu. Po niekoľkých kríženiach nakoniec zúžime sledovaný segment s QTL na menej ako 1 cM (30). Takáto oblasť už zahŕňa omnoho prijateľnejší počet kandidátnych génov pre skúmanie ich funkcií.

Pre jemnejšie mapovanie v asociačných štúdiách sa často používa už popísaná metóda väzbovej nerovnováhy (LD – linkage disequilibrium). Predovšetkým je výhodná pri hľadaní génov pre genetické poruchy (31). Metóda LD vyžaduje husté zaplnenie markermi v intervale. Je založená na detekovaní chromozómových oblastí, ktoré sú identické pôvodom (identical by descent – IBD). IBD sú alely na lokuse u potomkov, ktoré sú presnými kópiami alel prítomných už v nejakej predchádzajúcej generácii a neboli teda pozmenené žiadnou novou mutáciou, iba klasicky procesom pohlavného rozmnožovania prenesené do ďalšej populácie. LD teda vyhľadáva segmenty genómu, ktoré sú IBD medzi postihnutými jedincami, pretože tento región môže niesť gén pre danú poruchu (32). Pre meranie väzbovej nerovnováhy (LD) medzi QTL a markerom sa používa regresná analýza, ktorá odhadne vplyv markeru na kvantitatívny znak (33).

#### 4.1.3. Sekvenčná analýza

Dostatočne malý interval už možno podrobiť DNA sekvencovaniu, ktoré nám ukáže prípadné rozdiely v nukleotidových sekvenciách či už v kódujúcej alebo susednej oblasti. I minimálny interval môže obsahovať viacero génov a medzigenových sekvencií. Sekvenčná analýza nám ukáže varianty so zmenou jedného i niekoľkých nukleotidov, ktoré sa týkajú QTL. Tieto obmeny nachádzame vo vnútri jedného génu alebo vo viacerých funkčne príbuzných i nepříbuzných génov. Všetky kombinácie génov sú potom kandidátmi na identifikovanie a testovanie ich funkcie. Funkčný test býva založený na tzv. knock-in technológii (34), pri ktorej je transgén, čo je testovaný gén cielene pozmenený v oblasti sekvenčnej variability, inserovaný do vybraného lokusu namiesto pôvodne sledovaného genu. Pozoruje sa, či zmena nukleotidových variant vyvolá nejaký posun vo fenotypovej hodnote.

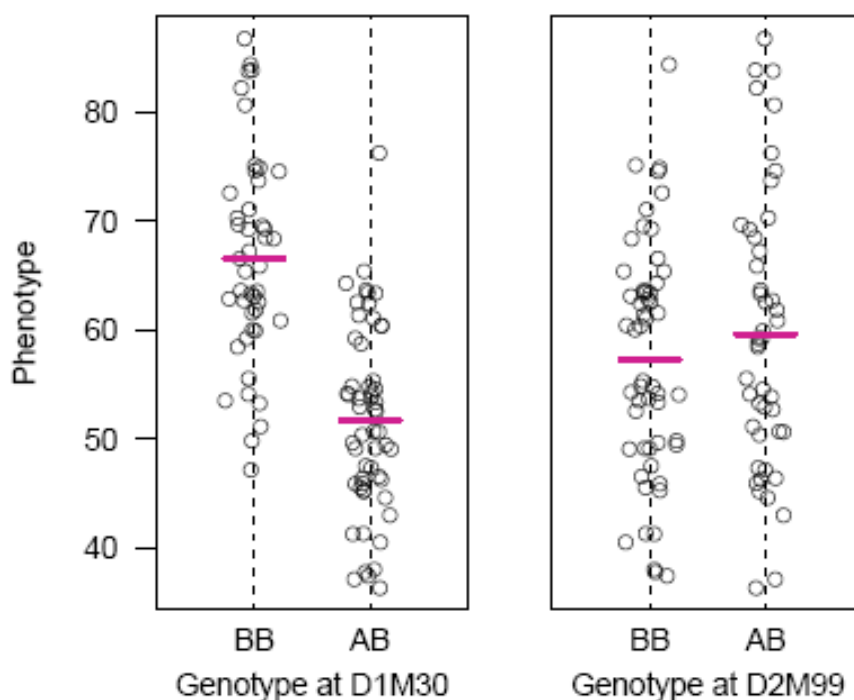
## 4.2. Metódy pre analýzu QTL

### 4.2.1 Analýza jednotlivých markerov

Asi najbežnejšou metódou pri detekcii QTL v susedstve je analýza jednotlivých markerov, ktorá študuje markery jeden po druhom. Potomkovia krížení sa u každého markeru rozdelia do skupín podľa svojich genotypov (13). U každej skupiny sa vypočíta stredná hodnota pre pozorovaný znak a hodnoty sa porovnávajú. Výhodní sú spätní kríženci dávajúci iba dve skupiny, keďže na každom markeri môžu mať len jeden z dvoch možných genotypov (AA alebo AB). Na zistenie štatistickej významnosti sa používa jednoduchý štatistický T-test (35) alebo ANOVA (Analysis of Variance- Analýza roptylu). Ak sa priemerné hodnoty znaku v skupinách pre daný marker výrazne líšia, je tento marker spojený s QTL a naopak ak sú približne rovnaké, uvažujeme, že nie je spojený s QTL. (obr. 3) Rozdiel v stredných hodnotách fenotypu nám odhaduje veľkosť efektu QTL.

Tiež je možné použiť regresnú metódu, ktorá si genotypy zakóduje na určité číselné hodnoty (napr. 0 a 1) a vynesie ich do grafu v závislosti od fenotypovej hodnoty. Preložením regresnej priamky sa dá odhadnúť podľa jej sklonu efekt QTL.

Tieto jednoduché metódy však vyradzujú zo štúdie jedincov s odlišným genotypom na markeri. Navyše môže byť umiestnenie QTL odhadnuté nepresne, iba na základe najvyšších rozdielov medzi skupinami u niektorých markerov. Následkom rekombinácie medzi markerom a QTL bude skutočný efekt QTL vyšší než vypočítaný. Účinnosť QTL detekcie býva znížená i širokým roložením markerov, pretože QTL sa môžu nachádzať vo väčšej vzdialenosti od markerov. Všetky tieto nevýhody podnietili rozvoj novších, presnejších metód.



Obr.3 : Príklad pre ANOVA test na markeroch D1M30 a D2M99. Bodky znamenajú hodnoty fenotypov.(Broman, 2001)

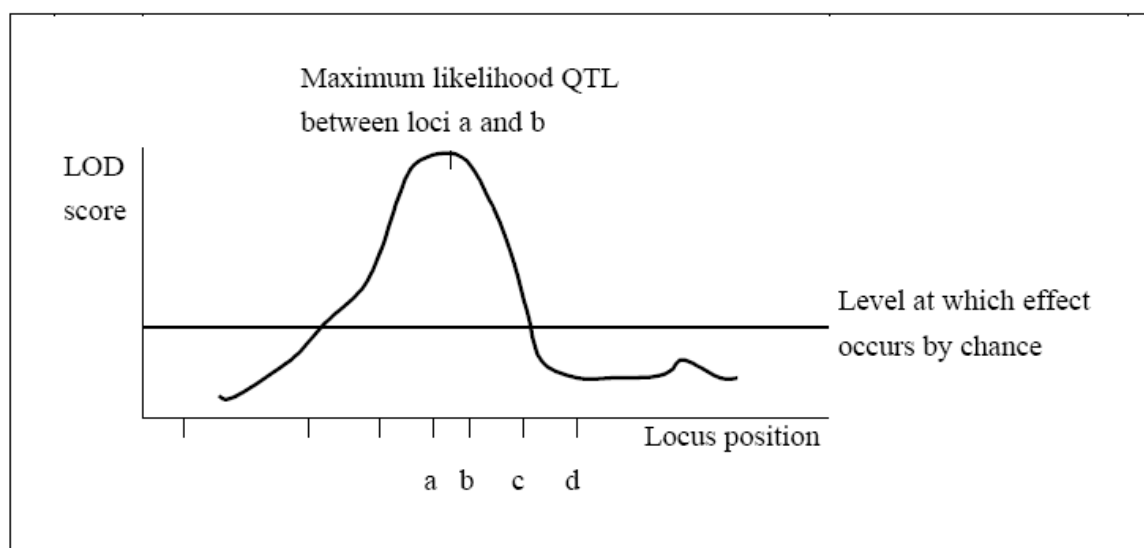
#### 4.2.2. Intervalové mapovanie

##### 4.2.2.1. Metóda maximálnej pravdepodobnosti

Snáď najobľúbenejšou metódou je intervalové mapovanie vyvinuté Landerom a Botsteinom (24). Je výsledkom kombinácie viacerých zlepšení v mapovaní QTL, ako prechodu na Bayesovo mapovanie s modelom pre početné QTL oproti staršiemu modelu s jedným QTL (36), začlenenie náhodných vplyvov do výpočtov (37) ako i zovšeobecnenie rozdielnych experimentálnych návrhov pre mapovanie (38). Intervalové mapovanie využíva genetickú mapu s molekulárnymi markermi a vyhľadáva v genóme QTL medzi dvoma susednými markerovými lokusmi. Vypočítava sa pravdepodobnosť prítomnosti určitého genotypu na predpokladanom QTL v závislosti od toho, aké genotypy majú jeho najbližšie susedné markery. Následne fenotypový efekt bude zmesou dvoch fenotypových distribúcií podľa genotypov markerov obklopujúcich QTL. Pre každú triedu markerov sa vypočíta odhad

maximálnej pravdepodobnosti, ktorý podľa pozorovaných údajov udáva najpravdepodobnejšiu hodnotu parametrov rozdelenia kvantitatívneho znaku. Prítomnosť QTL sa zisťuje na základe výpočtu LOD skóre (Likelihood of odds) (7). LOD skóre je vlastne logaritmom odhadu maximálnej pravdepodobnosti a ukazuje aká je pravdepodobnosť vzhľadom k pozorovaným dátam, že zvyšujú predpoklad prítomnosti QTL než jeho absencie. Teda ukazuje, aký silný je dôkaz o prítomnosti QTL, pričom čím vyššia je jeho hodnota, tým silnejší je dôkaz. LOD skóre sa vypočíta takto:

$$\text{LOD} = \frac{\text{pravdepodobnosť, že efekt sa vyskytne ak tam je prepojenie s QTL}}{\text{pravdepodobnosť že efekt sa vyskytne náhodou}}$$



Obr.4: Zjednodušený príklad pre intervalové mapovanie metódou odhadu maximálnej pravdepodobnosti rozdelenia LOD skórov, ktoré určuje umiestnenie QTL na mape .

#### 4.2.2.2. Regresná metóda

Mapovanie metódou maximálnej pravdepodobnosti je relatívne zložité a zdĺhavé pre počítačové výpočty, preto bola vytvorená jednoduchšia metóda pre intervalové mapovanie, metóda regresnej analýzy (39). V informácii o polohe a efekte QTL poskytuje rovnaké výsledky. Princíp spočíva v tom, že genotypy QTL sú určené ako pravdepodobnosti podľa vedľajších markerov a vynesené proti fenotypovým hodnotám.



Najväčšou výhodou intervalového mapovania je, že pri jeho správnom prevedení dovoľuje odhadnúť nekompletné genotypy markerov.

#### **4.2.3. Zložené intervalové mapovanie**

Prístup intervalového mapovania prináša určité obmedzenia, pretože pokiaľ sú dva QTL prepojené, kombinácia ich vplyvov spôsobí neobjektívnosť odhadov. Preto bola vyvinutá ešte komplexnejšia metóda a to zložené intervalové mapovanie (40). Spája v sebe metódu intervalového mapovania a mnohonásobnej regresie. Dá sa využiť pri mapovaní viacnásobných QTL, keď následkom ich rozchodu (pri kríženiach) môže byť určitý konkrétny QTL zaťažený chybou kvôli vplyvu zvyšných QTL. Ich účinok možno odstrániť vnesením regresie na markeroch ležiacich mimo intervalu so sledovaným QTL. Vedľajšie efekty sa teda neprejavia a toto mapovanie tak umožňuje väčšiu precíznosť.

#### **4.3. Počítačové programy pre QTL mapovanie**

V súčasnosti je dostupných viac ako 100 programov pre QTL mapovanie. Nižšie uvádzam najpoužívanejšie z nich.

##### **SAS**

Je program pre všeobecné štatistické analýzy ako ANOVA, T-test či GLM, dá sa použiť pri analýze jednotlivých markerov. Zisťuje spojitosť medzi genotypom markeru a fenotypom.

##### **MAPMARKER/EXP**

Je verejne dostupný softvér. Využíva údaje z experimentálnych populácií na zostavovanie genetických máp (41).

##### **MAPMARKER/QTL**

Analyzuje údaje z generácie F2 alebo spätných krížencov metódou intervalového mapovania (42).

##### **QTL Cartographer**

Taktiež je verejne dostupný. Umožňuje regresnú analýzu jednotlivých markerov, klasické i zložené intervalové mapovanie, vypracúva chýbajúce genotypové data na markeroch (43).

##### **MAPQTL**

Je program pre intervalové mapovanie s využitím metód pre modely s viacnásobnými QTL (44).

### **MAPMANAGER/QTX**

Tento program detekuje a lokalizuje QTL pomocou regresnej analýzy, jednoduchého i zložitého intervalového mapovania a zaznamenáva ich interakcie. Tiež poskytuje možnosť používať Haldanovou i Kosambiho metódou (45).

## **5. Záver**

Hľadanie génov pre kvantitatívne znaky i dnes vyžaduje zdĺhavé a zložitejšie postupy ako monogénna dedičnosť. Je ťažké nájsť všetky gény vysvetľujúce variabilitu určitého kvantitatívneho znaku, pretože vopred nevieme, koľko génov presne sa na prejave podieľa. Navyše niektoré z týchto génov majú len veľmi malý vplyv na zmenu hodnoty znaku a nezaznamenajú sa tak ako štatisticky významné. Každý ďalší objav QTL, resp. už konkrétneho génu ale znamená krok vpred a urýchľuje ďalšie bádania v tejto oblasti. Výrazné zjednodušenie prináša príprava kmeňov konsomických myší v našom laboratóriu. Oproti náročnému detekovaniu QTL v celom genóme u nich môžeme zamerať pozornosť iba na určitý chromozóm, ktorý je v danom kmeni substituovaný, ak sa tento kmeň spája s rozdielnou hodnotou sledovaného znaku oproti inbredným kmeňom. Existuje databáza, kde možno nájsť genotypy jedincov generácií vznikajúcich spätným krížením konsomikov s jedným rodičovským kmeňom a následným krížením v rámci generácie, ako i hodnoty pre niektoré znaky u každého jedinca. Tieto údaje poskytujú množstvo materiálu pre QTL mapovanie a nabádajú k zisťovaniu podstaty kvantitatívnych znakov, ktorých je mnoho a často sa jedná o dôležité znaky rozličných biologických systémov ako správanie, fyziológia, morfológia alebo náchylnosť ku všeobecným chorobám a poruchám. Následná identifikácia génov pre komplexné znaky odhaľuje genetickú príčinu ich povahy a dáva možnosť pokúšať sa o ovládanie zmeny ich prejavov.

## **Použité skratky**

**QTL** – quantitative trait loci

**PCR** – polymerase chain reaction

**SNP** – single nucleotide polymorphism

**DNA** – deoxyribonucleotid acid

**BAC** – bacterial artificial chromosome

**HS** – heterogenous stock

**LD** – linkage disequilibrium

**IBD** – identical by descent

**LOD** – likelihood of odds

## Prehľad použitej literatúry

1. Broman K.W., Review of statistical methods for QTL mapping in experimental crosses, *Lab Animal* 30: 44–52, 2001
2. McPeck M. S., From mouse to human: Fine mapping of quantitative trait loci in a model organism, *Proc Natl Acad Sci U S A.* ,97:12389-90, 2000
3. Collins, F. S., Guyer, M. S. & Charkravarti, A., Variations on a theme: cataloguing human DNA sequence variation, *Science* 278: 1580–1581, 1997
4. Choi H.-K., Kim D., Uhm T., Limpens E., Lim H., Mun J.-H., Kalo P., Penmetsa P.V., Seres A., Kulikova O., Roe B.A., Bisseling T., Kiss G.B., and Cook D.R., A Sequence-Based Genetic Map of *Medicago truncatula* and Comparison of Marker Colinearity with *M. sativa*, *Genetics*, 166: 1463-1502, 2004
5. Haldane J.B.S., The combination of linkage values and the calculation of distance between loci of linked factors , *J Genet.*, 8: 299–309, 1919
6. Touré A., Haussmann B.I.G., Jones N., Thomas H., and Ougham H., Construction of a genetic map, mapping of major genes, and QTL analysis
7. Doerge R. W., Weir B. S. and Zeng Z-B., Statistical issues in the search for genes affecting quantitative traits in experimental populations, *Statist. Sci.* 12: 195-219, 1997
8. Vos P., Hogers R., Bleeker M., Reijans T., Van De Lee T. et al., AFLP: a new technique for DNA fingerprinting, *Nucleic Acids Res.* 23: 4407–4414, 1995
9. Dakin E.E. and Avise J.C. , Microsatellite null alleles in parentage analysis, *Heredity* 93: 504 – 509, 2004
10. Paterson A. H., Lander E. S., Hewitt J. D., Peterson S., Lincoln S. E., Tanksley S. D., Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms  
*Nature* 335: 721 – 726, 1988
11. Lande R., Thompson R., Efficiency of marker-assisted selection in the improvement of quantitative traits, *Genetics.*, 124:743-56, 1990
12. Darvasi, Weinreb A., Minke V., Wellert J. I. and Soller M., Detecting Marker-QTL Linkage and Estimating QTL Gene Effect and Map Location Using a Saturated Genetic Map, *Genetics.* 134: 943-51, 1993

13. Mott R., Talbot C. J., Turri M.G., Collins A.C., and Flint J., A method for fine-mapping quantitative trait loci in outbred animal stocks, *Proc Natl Acad Sci U S A.*; 97,12649-54, 2000
14. Romanenko S.A., Perelman P.L., Serdukova N.A., Trifonov V.A., Biltueva L.S. et al, Reciprocal chromosome painting between three laboratory rodent species, *Mamm Genome.* 17:1183-92, 2006
15. Copeland N.G., Jenkins N.A., Gilbert D.J., Eppig J.T., Maltais L.J., Miller J.C., Dietrich W.F., Weaver A., Lincoln S.E., Steen R.G., et al., A genetic linkage map of the mouse: current applications and future prospects, *Science*, 262: 57-66, 1993
16. Dietrich W.F., Miller J., Steen R., et al., Comprehensive genetic map of the mouse genome. ,*Nature*, 380: 149–152, 1996
17. Nadeau J.H., Singer J.B., Matin A., Lander E.S., Analysing complex genetic traits with chromosome substitution strains, *Nat Genet* 24:221-5, 2000
18. Jansa P., Divina P., Forejt J., Construction and characterization of a genomic BAC library for the *Mus m. musculus* mouse subspecies (PWD/Ph inbred strain). *BMC Genomics.* 6:161, 2005
19. Legare M.E., Bartlett F.F.& Frankel W.N., A major effect QTL determined by multiple genes in epileptic EL mice. *Genome Res.* 10: 42-48, 2000,
20. Thoday J.M., Locations of polygenes, *Nature* 191: 368-370, 1961
21. Keightley P.D., Hardge T., May L., Bulfield G., A genetic map of quantitative trait loci for body weight in the mouse, *Genetics.* 142: 227-35, 1996 Jan;
22. Vallejo R.L., Bacon L. D., Liu H. C., Witter R.L., Groenen M.A., Hillel J., and Cheng H.H., Genetic mapping of quantitative trait loci affecting susceptibility to Marek's disease virus induced tumors in F2 intercross chickens. *Genetics.* 148: 349–360, 1998
23. Wolf N., Galecki A., Lipman R., Chen S., Smith-Wheelock M., Burke D., and Miller R Quantitative Trait Locus Mapping for Age-Related Cataract Severity and Synechia Prevalence Using Four-Way Cross Mice, *Invest Ophthalmol Vis Sci.* 45: 1922-9, 2004
24. Lander E.S., Botstein D., Mapping mendelian factors underlying quantitative traits using RFLP linkage maps, *Genetics.* 121: 185-99, 1989

25. Keightley P.D., Bulfield G., Detection of quantitative trait loci from frequency changes of marker alleles under selection. , *Genet Res.* 62:195–203, 1993
  
26. Mott R., Flint J., Simultaneous detection and fine mapping of quantitative trait loci in mice using heterogeneous stocks, *Genetics.*; 160: 1609-18, 2002
  
27. Crepieux S., Lebreton C., Servin B. and Charmet G., Quantitative trait loci (QTL) Detection in Multicross Inbred Designs“ Recovering QTL Identical-by-descent Status Information from marker data, *Genetics*, 168: 1737-1749, 2004
  
28. Glazier A.M., Nadeau J.H., Aitman T.J., Finding genes that underlie complex traits. *Science.*; 298:2345-9, 2002
  
29. Lander E. S., Kruglyak L., A Nonparametric Approach for Mapping Quantitative Trait Loci, *Genetics*, 139: 1421-1428, 1995
  
30. Ikeda A., Zheng Q. Y., Zuberi A. R., Johnson K. R., Naggert J. K. & Nishina P. M. Microtubule-associated protein 1A is a modifier of tubby hearing (moth1), *Nature Genetics* 30: 401 – 405, 2002
  
31. Hästbacka J., de la Chapelle A., Kaitila I., Sistonen P., Weaver A., Lander E., Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland., *Nat Genet.* 2: 204-11, 1992
  
32. Pritchard C., Cox D.R., Myers R.M., The end in sight for Huntington disease?, *Am J Hum Genet.* 49: 1-6, 1991
  
33. Meuwissen T. H. E. and Goddard M. E., Fine Mapping of Quantitative Trait Loci Using Linkage Disequilibria With Closely Linked Marker Loci, *Genetics*, 155, 421-430, 2000
  
34. Darvasi, Weinreb A., Minke V., Wellert J. I. and Soller M., Detecting Marker-QTL Linkage and Estimating QTL Gene Effect and MapLocation Using a Saturated Genetic Map, *Genetics.* 134: 943-51, 1993
  
35. Nebert D.W., Dalton T.P., Stuart G.W., Carvan M. J., Gene-swap knock-in" cassette in mice to study allelic differences in human genes, *Ann N Y Acad Sci.* 919:148-70, 2000
  
36. Zhang M., Montooth K. L., Wells M.T., Clark A.G. and Zhang D., Mapping Multiple Quantitative Trait Loci by Bayesian Classification, *Genetics*, 16: 2305-2318, 2005
  
37. Hoeschele I. and Van Raden P., Bayesian analysis of linkage between genetic markers and quantitative trait loci. II. Combining prior knowledge with experimental evidence , *Theor. Appl. Genet.* 85:946-952, 1993

38. Carbonell E. A., Gerig T. M., Balansard E. and Asins J., Interval mapping in the analysis of nonadditive quantitative trait loci, *Biometrics* 48:305-315, 1992
39. Haley C.S., Knott S.A., A simple regression method for mapping quantitative trait loci in line crosses using flanking markers, *Heredity*. 69: 315-24, 1992
40. Zeng Z.B., Precision Mapping of Quantitative Trait Loci, *Genetics* ;136: 1457–1468, 1994
41. Lander E.S. and Green P., Construction of multilocus genetic linkage maps in humans, *Proc Natl Acad Sci U S A.*; 84: 2363–2367, 1987
42. Lincoln S., Daly M. and Lander E., Mapping genes controlling quantitative traits with MAPMAKER/QTL 1.1, 2nd ed. Technical report, Whitehead Institute, Boston, MA., 1992
43. Basten C. J., Weir B. S. and Zeng Z.-B. ,QTL CARTOGRAPHER: A Reference Manual and Tutorial forQTL Mapping., Dept. Statistics, North Carolina State University,Raleigh, (1995{1996).
44. Van Ooijen J. W. and Maliepaard C. , MapQTL (tm)Version 3.0 Software for the Calculation of QTL Positions on Genetic Maps. DLO, Centre for Plant Breeding and Reproduction Research, (1996).
45. Manly K.F., Cudmore Jr. R.H., Meer J.M., Map Manager QTX, cross-platform software for genetic mapping, *Mammalian Genome* 12: 930-932, (2001)